Automatic resetting method intended for a geometric model of a scene over a picture of the scene, implementing device and programming medium

5   The present invention concerns an automatic resetting method, notably using electronic means, an implementing device as well as a programming medium, intended for resetting a geometric model of a scene over a picture of the scene. It finds applications in activities implementing the
10   processing of pictures and wherein one seeks to superimpose a visual model representing a scene according to a particular viewpoint and a particular vision angle on a real picture of the scene taken according to another viewpoint and/or angle.

The invention is more particularly intended for scenes
15   which exhibit references in the form of lines contrasted with respect to the remainder of the scene and, notably, of the sports grounds with reference marking and delineating lines. The resetting between the model and the picture enables to place an action unfolding on the scene in a significant context
20   provided by the model and linked with its structuration (location on the scene, knowledge of the usual actions of the location...). The invention enables therefore structuration of the picture. The pictures are notably video pictures.

Processes for analysing digitalised, video pictures or
25   others, enabling to extract automatically the characteristics of the picture, are known already. Such processes implement two approaches. The first, which is a general approach, is operational regardless of the type of picture processed. The second, which is a specialised approach, is adapted to the
30   type of the picture to be processed. With the first, the results obtained are relatively poor.

It is therefore desirable to specialise the analysing processes relative to the type of the picture. A particular type of picture has structuration elements which are particularly
35   interesting for this object and these are the broadcasting pictures of sportive events taking place on particular grounds

having reference marking and delineating lines. Indeed, besides the fact the rules of the game are generally quite simple, which enables structuration of the match and easy recognition of the characteristic actions, the location whereon

5   the action takes place, is fixed and its spatial structure is known a priori relative to references which are relatively simple to detect, i.e. lines or curves. By way of example, running or racing circuits may be mentioned (athletics or sports cars), European or American football pitches,

10   basketball pitches, and tennis grounds. For the latter, it is generally possible to known in advance the sequence of operations of a game, i.e. the possible chaining of the game phases (temporal action structuration), the universal model of the ground with the accurate dimensions of the different lines

15   (spatial structuration), the number of players, etc.

Several solutions have already been suggested for solving the problem associated with the resetting of a sports ground model on pictures. Three examples may be mentioned, regarding three different types of sport, i.e.

20   tennis, football (soccer), the American football.

For tennis, in the article of G. Sudhir, J. Lee and A. Jain, entitled « Automatic classification of tennis video for high level content-based retrieval » Technical Report, August 1997, The Hong-Kong University of science and technology,

25   one endeavours to find thee perpendicular lines in the picture (a service square, for example) in order to calculate the position of the other lines knowing the theoretic form of a tennis ground. The first step regarding the recognition of three lines, is carried out by a line-tracking algorithm

30   restricted by advance knowledge of the seeking direction (horizontal to the right, then vertical upwards, and finally horizontal to the left). This algorithm is initialised by a point selected heuristically in the centre of the picture. The major defects of this approach are the lack of robustness in the

35   positioning of the starting point, also in case of the absence

of one of the three perpendicular lines (noise in the picture, lines partially effaced) and the lack of verification of the adequation between all the resetted lines and the lines in the picture (only the three base lines are tested). Finally, the process suggested leads to rather slow an algorithm, little appropriate to real-time processing.

As regards football, in the article of Y. Gong T.S. Lim, and H.C. Chua, entitled « Automatic Parsing of TV Soccer Programs » IEEE International Conference on Multimedia Computing and Systems, May, 1995, pp. 167-174, the following steps are implemented: Contour detection by a Laplace-Gauss filter; Filtering of contour information by using the white colour of the lines; Form recognition (ellipse, triangle, rectangle...) giving a number of primitives and analysis of the spatial relations between primitives enabling to identify the point of the ground where the action the game is taking place. Such process is particularly adapted to the football pitch by reason of the heterogeneity of the primitives sought (kick-off area, penalty area, goal...). It proves more difficult to apply to a ground model which exhibits a large symmetry such as a tennis ground. Moreover, it does not provide any resetting of the ground properly speaking, but rather recognition of the position on the ground (close to goals, in the centre...).

Finally, as regards American football, in the thesis of S. Intille, entitled « Visual Recognition of Multi-Agent Action ». Phd Thesis, MIT, September 1999, the characteristics of the American football ground are used in order to recognise the point when the game action is taking place. To do so, one uses the markings on the ground. These are composed of figures and of lines distributed every « 10 yards ». These pieces of information are collected within a theoretic ground model. The method suggested then consists in matching n (n>=4) points of the picture with n points of the theoretic ground. To do so, a line detection algorithm based on Cany-

Deriche filtering is used. The intersections of the different straight lines found form a collection of particular points serving for resetting with respect to the theoretic model. The initial resetting is performed manually on the first picture

5 while associating 4 points identified in the picture with their counterpart in the theoretic model. For the following pictures, an algorithm for compensation of the dominant movement enables to track the matching points throughout the sequence. The shortcomings of such method are mainly the

10 use of manual initialisation, the sensitivity of the line detection algorithm and the difficulty of adaptation to a more complex ground model which does not exhibit any equivalent ground markings.

The present invention suggests an alternate method

15 which does not resort to manual initialisation of the resetting algorithm for each video sequence processed. It is moreover robust to the problem associated with contour detection, which is not the case of the methods described previously. Within the framework of the invention, the terms ground and

20 scene are considered as equivalent.

Thus, the invention concerns, an automatic resetting method using electronic means intended for a geometric model of a scene over a picture of the scene, the model and the picture of the scene being stored in the memory of an

25 electronic device in the form of pixel matrices, the scene including fixed references with respect to the remainder of the scene, whereas the references may be specifically detected within the matrices, the picture being taken by a camera arranged in a given zone with respect to the ground

30 in a location of the zone and according to a shot angle determined relative to the scene, the electronic means comparing the picture with the model having been adjusted in perspective by homography for superimposition of the references.

According to the invention, the electronic device calculates a fine homography function H_f for resetting into three main steps:

- a first preliminary phase of determination of an average resetting homography consisting in determining an average homography function H_m applicable to the model with average adjustment over a sample of pictures of the scene taken previously,

- a second, rough resetting phase consisting after application of the average homography function H_m to the model in determining a rough homography function H_g,

- a third, fine resetting phase consisting after application of the rough homography function H_g to the model in determining a fine homography function H_f.

It should be noted as of now that, as can be seen later on, the shot location and/or the shot angle may evolve from one picture to the other inasmuch as the model remains partially visible in the picture (the visibility limit criterion will be defined at a later stage). In diverse implementation modes of the invention, whereas the following means may be used alone or in combinations according to all technical possibilities, are employed :

- the scene possesses reference delineating or marking lines, i.e. at least 4 reference lines, non-parallel 3 by 3,

- the reference lines are reduced to points and the scene possesses at least 4 reference points, non-aligned 3 by 3, ··

- in the preliminary phase of determination of an average resetting homography, at least one sample picture is selected among a collection of pictures taken of the given location, the references on the sample picture(s) are detected and one calculates an average homographic function H_m enabling superimposition between the model subjected to the average homographic function and the sample picture(s), superimposition being reached for least error square minimization of the distance between reference points of

sample picture(s) and the model subjected to the average homographic function,

- in the second, rough resetting phase:

    - in a first step, an extraction process is applied to the picture enabling, according to detection criteria, to detect in the picture matrix of the pixels liable to represent references of the scene (but hence seen according to the shot location and the shot angle) and to form a first binary matrix M_{rh} of picture reference including horizontal contour points (also called vertical gradient points) and a second binary matrix M_{rv} of picture reference including vertical contour points (also called horizontal gradient points),

    - in a second step, one calculates for each horizontal reference binary matrix M_{rh}, respectively vertical reference binary matrix M_{rv}, a horizontal reference distance matrix M_{dh}, respectively a vertical reference distance matrix M_{dv}, including for each element of the matrix the distance value with respect to the closest reference according to the vertical line, respectively the horizontal line,

    for the horizontal reference distance matrix M_{dh} each element of said matrix specifying the distance in number of pixels relative to the reference line along a vertical axis, the distance values on the reference line and those of a column without any reference line pixel being nil, the distance values along the vertical line increasing in absolute value as the element moves away relative to the reference line, the distance values of the elements being of opposite signs on both sides of the reference line,

    for the vertical reference distance matrix M_{dv} each element of said matrix specifying the distance in number of pixels relative to the reference line along a horizontal axis, the distance values on the reference

line and those of a line without reference line pixel being nil, the distance values along the horizontal line increasing in absolute value as the element moves away relative to the reference line, of the elements being of opposite signs on both sides of the reference line,

- in a third step, all the reference lines of the model are applied the average homographic function H_m in order to produce a binary average adjusted matrix M_{am} which is compared with the vertical M_{dv}, respectively horizontal M_{dh} reference distance matrices, for pixel matching purposes,

with for each pixel $p(i,j)$ of the average adjusted matrix derived from a resetted pixel of the model belonging to a vertical reference line and positioned at the line i and at the column j of the average adjusted matrix M_{am}, the allocation of a corresponding pixel obtained by adding the value v in i and j of the vertical reference matrix M_{rv} to the value j, and matching the pixels $((i,j), (i,j+v))$,

with for each pixel $p(i,j)$ of the average adjusted matrix derived from a resetted pixel of the model belonging to a horizontal reference line and positioned at the line i and at the column j of the average adjusted matrix M_{am}, the allocation of a corresponding pixel obtained by adding the value v in i and j of the horizontal reference matrix M_{rh} to the value i, and matching the pixels $((i,j), (i+v,j))$,

a homography function H_{opt} is then calculated by regression with minimisation of the medial of the square of the distance between pairs of matched pixels, the calculation being carried out over n collections of four pairs of matched pixels,

- in a fourth step, one identifies the pairs of pixels corresponding to non-aberrant matches,

- in a fifth step, H_{opt} is adjusted by least square regression calculation over all the non-aberrant pixel pairs in order to produce the rough homography H_g,

- in the binary average adjusted matrix M_{am}, the pixels take
5   on the value 1 if they correspond to a reference pixel of the resetted model and 0 if not,

- in the fourth step of the second, rough resetting step, a pair of pixels corresponds to a non-aberrant match, if, for the pixel of the average adjusted matrix M_{am} of the match in
10   question, the distance between the pixel matched by using the reference matrices M_{rh}, M_{rv}, and that obtained by the homography H_{opt} is smaller than or equal to a preset threshold,

- the reference detection criteria are chosen individually or in
15   combination among:

   - a specific colour of the reference with respect to the remainder of the scene,
   - a specific tone of the reference with respect to the remainder of the scene,
20   - a specific grey level of the reference with respect to the remainder of the scene,
   - a specific shape of the reference, notably a line, an angle .between two lines ·crossing each other, a parallelism between two lines,
25   - a specific orientation of the reference,
   - a line closest and parallel to an edge of the picture matrix,

- the extraction process comprises a preliminary Cany-Deriche filtering step of the picture in order to obtain a
30   gradient picture and the processing resumes with the gradient picture,

- in the third, fine resetting phase the rough homography H_g is applied to the model and the result is compared with both horizontal and vertical distance matrices with adjustment of

the homography by a so-called Powel alternate single-dimension iterative minimisation method,

- the matrix of the model is a binary matrix whereof all the pixels regarding the references have a first value and the other pixels a second value in order to dispense with the detection of the references in said matrix of the model when implementing the phases and steps of the method,

- the pictures evolve with time according to of the sequences corresponding to different shot locations and/or angles and the electronic device comprises means enabling moreover to determine during the first, average resetting preliminary phase as many average homography functions H_m as there are different shot locations and angles,

- the phases and steps are implemented in the electronic means which are programmable logic units with a programme and the programmable logic comprises a microprocessor or a digital signal processor (DSP) and, preferably, of the general-purpose or dedicated microcomputer type,

- the scene is a sports ground including references in the form of delineating lines, notably a European or American "football" pitch or a tennis ground,

- electronic means are implemented which are wired logic units,

- the wired logic unit comprises at least one integrated circuit,

- electronic means are implemented which are programmable logic units with a programme,

- the programmable logic unit comprises a microprocessor or a digital signal processor (DSP) and are, preferably, of the general-purpose or dedicated microcomputer type.

The invention also concerns an automatic resetting device using electronic means intended for a geometric model of a scene over a picture of the scene, the model and the picture of the scene being stored in the memory of an electronic device in the form of pixel matrices, the scene including fixed references with respect to the remainder of

the scene, whereas the references may be specifically detected within the matrices, the picture being taken by a camera arranged in a given zone with respect to the ground in a location of the zone and according to a shot angle

5   determined relative to the scene, the electronic means comparing the picture with the model having been adjusted in perspective by homography for superimposition of the references.

According to this invention, the device comprises means

10   enabling to calculate a fine homography function H_f for resetting into three main phases:

-     a first preliminary phase of determination of an average resetting homography consisting in determining an average homography function H_m applicable to the model with average

15   adjustment over a sample of pictures of the scene taken previously,

-     a second, rough resetting phase consisting after application of the average homography function H_m to the model in determining a rough homography function H_g,

20   -   a third, fine resetting phase consisting after application of the rough homography function H_g to the model in determining a fine homography function H_f.

The device of the invention further comprises means enabling the execution of the method listed previously and of

25   all its variations, individually or according to all their combinations.

In a variation of the device, the electronic means are of the general-purpose or dedicated microcomputer type.

The invention also concerns an information storage

30   medium including a programme intended for operating the former device.

The invention finally concerns an information storage medium including a programme intended for operating the former device and at least according to one of the

35   methodological modalities among to all the modalities,

including those resulting from any possible combinations, listed previously.

The present invention will now be exemplified by the following description, without being limited thereto, and in

5   relation with:

Figure 1 which represents an example of device for implementation of the invention and,

Figure 2 which represents data generated for a rough resetting step with contour algebraic gradient and distance

10   cards.

The invention is now explained while taking as an example the resetting of a model of tennis ground over a picture coming from a game video sequence over such a ground, but which is generally taken according to another

15   viewpoint than that of the model. The tennis ground, advantageously, has dimensions which are perfectly known and reference lines perfectly defined. The object is the resetting of all the pieces of spatial information extracted from the video pictures, for example position and trajectory of

20   the players or of the ball, relative to a common referential which is the model. The resetting enables to define a transformation which may be then used for all the elements of the picture. It should be noted that according to what wants to be transformed, the model or a picture, one will use the

25   direct transformation or its reverse. This enables, in later phases, not covered in this application, to identify the phases of the game (service, volley ...).

In this example, one uses a number of hypotheses which are that the shots of the video pictures are made from

30   a high location behind the smaller side of the ground and that the major portion of the ground is visible on the pictures. However, the invention is applicable to pictures taken from another view point, notably on the larger sides. Moreover, one assumes that the ground lines are white (the invention is

35   however adaptable to any colour of line which may be

extracted from a picture of the ground). Finally, one assumes that the playing surface and hence the colour of the ground is not known in advance in order to be able to analyse the pictures of events on clay or grass tennis grounds and, also,

5   that the positioning of the players is a random one.

Building up on these hypotheses, one has determined a method which may be transposed into algorithmic form so that it may be implemented automatically in electronic means, notably a microprocessor or digital signal processor-based

10   computer system. The method uses the lines of the ground and more particularly the contour lines in order to reduce the quantity of calculations necessary. However, the invention may be applied to all the reference delineating and marking lines of the ground under the condition there are at least 4

15   reference lines, non-parallel 3 by 3 or 4 reference-points non-aligned 3 by 3.

With the method suggested, a piece of electronic equipment enables to deform automatically the theoretic ground model represented by delineating lines in order to

20   match as well as possible the resetted lines with the actual lines of the court which appear on the video pictures. In the computer equipment, the pictures as well of the model as the actual ones of the video, are in digital format and are stored in lines x columns matrices for the calculations. Preferably

25   the model corresponds to a binary picture of the scene (court) wherein the reference lines have a different value from of the remainder of the scene. Preferably, the picture matrices, the model matrices and those calculated have the same size in order to simplify the calculations and to avoid

30   needing to take into account a reduction factor or an enlargement factor. However, the invention may be applied in its principle to model and picture matrices of different sizes.

An example of device for implementation of the invention is represented on Figure 1 and, preferably, the

35   invention is implemented with programmable electronic

circuits, notably a dedicated microcomputer or data-processing equipment and one has represented on that figure data-processing means. A video camera 3 takes a shot of a scene 2 which is here a tennis ground. The shot including a

5   sequence of pictures is transmitted in the form of video data 1 to a microcomputer 4 enabling to perform the operations in relation with the invention and to store at least the picture under processing. In a circular box on the right of Figure 1 and in relation by an arrow with the microcomputer 4 one has

10   also represented a model 7 of tennis ground to indicate that the microcomputer also stores a representation of the model of the scene. The direct link between the camera 3 and the microcomputer enables direct processing of the video flux 1 which may be stored thereon. However, one has also

15   represented as a dotted line, a video link 1' between the camera 3 and a means 6 for storing the video flux for its first part and between the storage means 6 and the microcomputer 4 for its second part, in order to show that the invention may also apply to pre-recorded video. The storage

20   means 6 is represented in the form of a server, but it is also possible to use analogue storage means. However, it should be well understood that automated processing is carried out in a piece of equipment implementing logic/digital calculations, micro-processor or digital signal processor

25   (DSP), and that, if an analogue video signal is transmitted, an analogue/digital conversion is carried out before automated processing. Preferably, the video flux is a flux of digital data.

      One also understands that the term microcomputer may

30   cover any electronic computer equipment compatible and possibly dedicated of the graphic workstation type. Alternately, the microcomputer may be replaced with a wired circuit specifically realised to conduct the operations in relation with the invention. The wired circuit (one or several

35   integrated circuits) may possibly be arranged on an

electronic card within a microcomputer. Finally, the raw video data leading to particularly important data flux, the invention may also operate on compressed video data, whereas the pictures are decompressed during processing, or the
5   processing is suited to the type of compression used. In particular certain compression systems enable to suppress redundant data in a picture and a small quantity of data provides information regarding the homogeneity or the complexity of said picture, which may also enable the
10   selection of the shots.

      Resetting is carried out globally with a single model of deformation for the whole ground according to a particular shot location of the picture and inasmuch as at least 4 of the references associated with the model are visible. In case
15   when the shot location has been modified, a new model of deformation should be implemented (the average initial deformation model should be changed which implies modification of the results of the later steps). In the case of sequences which alternate shots in different locations, the
20   equipment may be provided beforehand with information on the shot location and one uses the corresponding average resetting function (the method implements a priory step of determination of an average resetting function determined on the basis of a sample of pictures taken in a particular point)
25   or, iterative tests are conducted with several average resetting functions (each corresponding to a particular point) looking for the resetting which is closest according to a distance criterion between resetted model and picture and one uses the average resetting function in question for the
30   remainder. It should be noted once more that, at a later stage, and outside the framework of the present invention which concerns more particularly the resetting between a model and a picture, once the different positions of the players or of the ball have been calculated on the actual

picture, they may also be compensated for via the resetting function or its reverse.

The principles at the base of the invention will now be explained, by considering a ground model which corresponds
5   to a ground seen from above, substantially in its centre and symmetrically. The invention may however be implemented with a ground model which corresponds to a different view. Nevertheless, one chooses preferably a view of the model which simplifies the calculations and, especially, the later
10   steps for positioning the elements of the picture.

A function enabling to deform this theoretic ground or model including references crossing each other at right angle, i.e. the model is considered as seen from above substantially in its centre, is therefore sought. One knows that the same
15   ground seen through a camera positioned laterally will exhibit on picture, reference lines in perspective, whereas the vanishing lines are not parallel, contrary to the same lines of the model. The type of projection to used to deform the model and to superimpose said model onto the picture is known and
20   it corresponds to a perspective projection function (the non-linear deformations associated with optical imperfections of the camera are neglected). Under this hypothesis of pure perspective projection, one knows that there exists an exact relation enabling to transform a plane, that of the model, into
25   its projection. This function is the eight-parameter homographic function. Although it is non-linear in Cartesian coordinates, the passage into homogeneous coordinates enables to find linearity between a point of the model and its projection in the picture.

30   The principle of this transformation ought to be reminded at this point. Let there be $\tilde{p}(x,y,t)$ a point 2D expressed in homogeneous coordinates (the case when t is nil corresponds to a point at infinity in the direction (x,y)). This very point expressed in the Cartesian space will have as
35   coordinates $p(\frac{x}{t},\frac{y}{t})$ (a point at infinity cannot be expressed in

Cartesian coordinates). In homogeneous coordinates, the homographic transformation is expressed in matrix form, via the matrix H(3×3) defined within one multiplicative coefficient (it possesses 8 independent coefficients). Whatever the homogeneous point $\tilde{p}^l_{th}$ expressed in the reference attached to the theoretic ground and $\tilde{p}^i$ its counterpart in the picture, one has $\tilde{p}^i = \lambda.H.\tilde{p}^i_{th}$ with λ as a non-zero scalar.

The resetting consists therefore in identifying the homography enabling to reset the theoretic ground, the model, on the actual picture. This type of identification is based on an adjustment iterative calculation which comprises a stopping condition based upon a criterion of quality. Normally, this quality criterion should be based upon the average distance between resetted reference lines and actual reference lines. However, the positioning of the actual lines is not known in advance. Consequently, the quality criterion which is used is a distance criterion to be minimised. Such criterion D(I,H), depending on the picture I and on the homography H, is defined as the integral along the resetted reference lines, of the distance between a point of a resetted reference line and the closest contour point. Its symbolic expression is as follows:

$$D(I,H) = \oint_T d_c(I,H.s).ds$$ where d_c is the Euclidian distance of the

point s resetted by the homography H at the closest contour point in the picture I.

However, the deformation to be applied to the theoretic model is very important and because of its nature highly non-linear in Cartesian coordinates, the homographic transformation is relatively unstable, small variations on the parameters of the third line of the homographic matrix causing very high variations in the position of the resetted points, which is not favourable to efficient automated calculation.

Consequently, the resetting method according to the invention will be carried out in three phases enabling to

switch from an approximate initial resetting to a fine-tuned final resetting. The three phases to be conducted are, firstly, an average resetting followed, secondly, by a rough resetting then, thirdly, by a fine resetting. The object of the invention

5   is, by using three resetting phases, to guarantee the production of a satisfactory solution with reduced cost (importance of the calculations) relative to the technique based upon the mathematical transformation in projection presented previously by way of reminder. The average

10   resetting phase uses the fact that the pictures exhibit close spatial characteristics. The rough resetting phase consists in cutting the scene (ground) into vertical and horizontal lines. The fine resetting phase is based upon a minimisation diagram whereof the rapid convergence towards satisfactory

15   minimum is guaranteed by the previous resetting phases.

    The resetting method of the invention may be explained in the form of an algorithm:

    1. Calculate over a representative set of pictures (sample), the average resetting homography H_m

20   2. Conduct rough picture resetting between the pixels of the resetted model by H_m and their counterparts obtained by the distance cards by:

    a. Calculation of the cards of horizontal and vertical gradients and of the cards of horizontal and

25   vertical distances with respect to their contour points (a contour point being a point whereof the gradient value is greater than or equal to a fixed threshold)

    b. Assessment of all the counterparts (couples) of

30   the points resetted by H_m on the basis of the distance cards,

    c. Robust calculation of the rough homography H_g on the basis of all the couples found (robust calculation consists in taking into account only a

35   portion of the couples, those which satisfy a

quality criterion regarding the matching process: the couples corresponding to non-aberrant matches)

3.   Conducting the fine resetting by using H_g by: fine-tuning of the parameters of H_g to minimise the actual contour distances on the picture with those of the theoretical ground (the model) resetted by H_g.

These three phases will now be described in detail:

a) the average resetting:

This first phase is carried out beforehand and, in case when the shot locations may be different, once at least for any shot location. It should be noted that in case when the shot angle may also evolve significantly, for example further to a « travelling » motion or to a rotation of the camera, said step may be carried out on the extreme angles and/or on intermediate shot directions (angular sectors).

A sample representative of pictures of the tennis sequences available for a given shot location (and, possibly a given shot angle) is determined and an average homographic function H_m, via least error square minimization of the distance between projected reference points and actual reference points, is calculated. Preferably, this step is carried out manually, an operating matching manually the visible reference angles in the pictures and the reference angles of the model (least square calculation of the average homography being carried out over all the matches thus obtained). However, such operation may also be conducted semi-automatically, an automaton adjusting the lines roughly and a human operator fine-tuning the adjustment to generate finally the average homographic function H_m. Conversely, this may be the human operator who adjusts roughly and the automaton which fine-tunes the adjustment to generate finally the average homographic function H_m.

b) the rough resetting

During this phase, two contour algebraic distance cards or matrices are calculated, one for vertical and one for horizontal. The term contour corresponds to reference lines of the ground. To this end and as represented on Figure 2, on
5 the basis of the original picture 8, pictures of vertical and horizontal gradients are calculated using Cany-Deriche filtering, then thresholding and binarisation relative to detection criteria in order to keep, in this example, only the highly contrasted points, i.e. those which correspond to the
10 reference lines, in order to generate two vertical contour 10 and horizontal 9 cards or matrices, respectively. If necessary, one may refer to the article Deriche, R., « *Optimal Edge Detection Using Recursive Filtering* », Proc. First Conf. on Computer Vision, London, June, 1987 as regards filtering. It
15 should be noted that on Figure 2, the contour cards 9 and 10 have been filtered moreover in order to generate at the outcome only the highly contrasted points with a grey level higher than a given threshold, whereas at this stage the points belonging to the reference lines are considered as
20 white. It should be noted that one may also take the colour into account to select contour lines or any other specific detection criterion of such contours in the picture (alignment of points, contrast, colour, crossing lines...). Alternately, in case when the reference lines were detectable simply, one
25 may use directly the picture and apply reference line detection criteria without going through the calculation of a gradient. These criteria may be a specific colour of line for example. It should be noted, finally, that it is possible to implement complementary steps enabling to improve the
30 quality of the reference lines detected, notably by expansion, erosion operations ... over the matrices. This enables for example to gather two portions of the same line which was cut either by the presence of a player in the axis, or by the covering of the ground pushed over the line by a slipping
35 payer (clay ground).

On the basis of two contour cards 9 and 10 thus calculated, one determines the algebraic distance cards with the closest horizontal 11 and vertical 12 contours. To this end, two horizontal (respectively vertical) scans of the

5   vertical 10 (respectively horizontal 11) contour card are carried out by allocating to each pixel the value of the Euclidian distance at the closest contour point (belonging therefore to a reference) on the line (respectively the column) scanned. This distance is negative before the contour point

10   on the line (respectively the column) scanned.

This phase of the method may be explained in the form of an algorithm while considering:

- Gh and Gv the cards of horizontal and vertical gradients corresponding to line x column matrices

15   indexed in i,j or in p (e.g. Gh(i,j): value of Gh at the point of coordinates i,j) (e.g. Gv(p): value of Gv at the point p)

- Dh and Dv the cards of horizontal and vertical distances corresponding to of the line x column

20   matrices indexed in i,j or in p (e.g. Dh(i,j): value of Dh at the point of coordinates i,j) (e.g. Dv(p): value of Dv at the point p)

- I(p) the intensity of the picture I at the point p

1. Calculation of the cards of horizontal Gh and vertical Gv

25   gradients

2. Binarisation of the gradient cards:

   a. Horizontal, for any point p:

      i. If (I(p)>threshold1) && (Gh(p)> threshold2) then Gh(p)=1

30   ii. If not Gh(p)=0

   b. Vertical, for any point p:

      i. If (I(p)>threshold1) && (Gv(p)> threshold2) then Gv(p)=1

      ii. If not Gv(p)=0

35   3. Calculation of the distances Dh and Dv:

a. Initialise Dh with the value (number of lines +1), i.e. $Dh(p) = Nblines + 1 \ \forall p \in Dh$

b. Initialise Dv with the value (number of columns +1), i.e. $Dv(p) = Nbcolumns + 1 \ \forall p \in Dv$

c. Calculation of Dh:

   i. For each column j:
- $d = -1$;
- n = number of lines
- For (i=0 to n-1)
  - a. if $Gh(i,j) = 1$ then $d = 0$
  - b. if $d \ != -1$ then
    - i. $Dh(i,j) = d$
    - ii. $d = d + 1$

   ii. For each column j:
- $d = -1$;
- n = number of lines
- For (i=n-1 to 0)
  - a. if $Gh(i,j) = 1$ then $d = 0$
  - b. if $(d \ != -1) \ \&\& \ (Dh(i,j) > d)$ then
    - i. $Dh(i,j) = -d$
    - ii. $d = d + 1$

d. Calculation of Dv:

   i. For each line i:
- $d = -1$;
- n = number of columns
- For (j=0 to n-1)
  - a. if $Gv(i,j) = 1$ then $d = 0$
  - b. if $d \ != -1$ then
    - i. $Dv(i,j) = d$
    - ii. $d = d + 1$

   ii. For each line i:
- $d = -1$;
- n = number of columns
- For (j=n to 0)
  - a. if $Gv(i,j) = 1$ then $d = 0$

b. if (d !=-1) && (Dv(i,j) > d) then
     i. Dv(i,j)=-d
     ii. d=d+1

Thanks to the distance cards 11 and 12 thus calculated, one will determine a rough homography H_g by a trial and error method with minimisation of the criterion $D(I,H)$. In this view, one applies the average homographic function H_m to the model to form an average adjusted model in the form of a matrix of the average adjusted model. The matrix of the average adjusted model and the card of vertical, respectively horizontal, distances are travelled in parallel, in one case horizontally and in the other vertically and the contour points of the picture are matched with their counterparts of the average adjusted model. The points for the picture are obtained on the basis of the card of vertical (respectively horizontal) distances. Two scans are carried out, a horizontal scan and a vertical scan. Thus, if $p(x,y)$ is a point of horizontal (respectively vertical) line of the adjusted model and d the value in (x,y) of the card of horizontal (respectively vertical) distances, then the counterpart of p in the picture will be the point of coordinates $(x,y-d)$ (respectively $(x-d,y)$).

The travel of all the matrices (distance cards and matrix of the average adjusted model) provides with a collection including a large number of elements of pairs of matching points which may however contain mistakably matches pairs of points (for example, pixels of the adjusted model non visible in the picture will be matched with the closest points with highest gradient (or outside the picture, failing any gradient points on the line or the column in question. In all cases, a match is provided with the closest point.

All these matches or pairs will now be used for calculating the new matrix of rough homography H_g transforming the theoretic ground points into points belonging to the contour cards. The technique used to this end is not based upon the least squares. Indeed, the card of the contour

points, regardless whether horizontal or vertical, often suffers from a significant noise. Certain matches may therefore be quite erroneous and direct least square assessment might lend too much importance to aberrant pairs which might,

5   because of the instability of the homographic model, cause divergence of the resetted model with respect to the actual positioning of the ground in the picture.

One prefers therefore to use a robust method of assessment of the parameters of the rough homography

10   function H_g. The object is to separate the matching pairs which are suitably matched from the aberrant matches. There exist several of families of robust assessment techniques. According to a preferred embodiment of the invention, the parameters of the rough homography H_g are calculated in

15   order to meet the least medial square criterion. The calculation method, presented briefly here, is described thoroughly at paragraph 3 of the article of P. Meer, D. Mintz and A. Rosenfeld « Robust Regression Methods for Computer Vision : A Review », published in International Journal of

20   Computer Vision, volume 6 n° 1, 1991, pages 59 to 70, to which may be referred.

According to such method, if one considers H the space of the parameters of the homography, E the collection of the matching pairs (called samples) and $c(p_{th}, p_r)$ a point-pixel

25   pair composed of a pixel of the average adjusted model p_{th} and of a matching contour point p_r in the picture, the least medial square method tends to minimise, in the space H, the medial of the residues calculated on E. The residue in the present case, corresponds to the distance in a pair between

30   the reference point resulting from the application of the current homography function to the model and the matching contour point (reference) on the picture. The homography H_{opt} solution of the problem is the homography minimising such medial:

35     $$H_{opt} = \min_{H_i \in H} (\operatorname{med}_{c \in E} d(H_i.p_{th}, p_r)^2)$$

where d is the Euclidian distance between two points.

The solution to the problem of non-linear minimisation calls for iterative search by regression of the homography H_i which minimises the medial of the residues among all the possible homographies. Preferably, in order to limit the calculations, one limits the search to a finite collection of n homographies, defined by n collections of four pairs (or couples) of points taken randomly in E. In implementation variations, one may use eight pairs, possibly sixteen pairs or more, according to the power of calculation available and/or the accuracy requested. For each of the n homographies, one calculates and one sorts the squares from the residues in order to identify the medial square residue. The resulting homography is assessed as that which provides the smallest medial square residue.

Selecting the homography on the single medial square residue, instead of over all the residues, confers to the regression process its robust character. Indeed, it enables not to take into account the residues of extreme values, liable to match aberrant pairs of points and henceforth to distort the regression.

It may be demonstrated statistically. By way of example, in case when eight pairs are used, if P=0.999 the probability that at least one of the n collections of eight pairs does not contain any aberrant couples and supposedly 50% of the data may be false; the number of draws n to be conducted to meet the probability P is then 1765. If the proportion of aberrant samples is smaller than 50%, supposedly, a collection which does not include any aberrant samples provides a resetted model in better keeping with the collection E, thereby showing a medial square residue which is smaller than any other collection including at least one aberrant sample. It is then almost sure that the homography finally obtained is defined by a collection of eight non-aberrant pairs, which guarantees the robustness of the method.

The homography H_{opt} obtained by this regression calculation is now used to identify the aberrant pairs and it is applied to the model to form a new optimal adjusted model. In this view, one calculates the standard deviation $\hat{\sigma}$ of the absolute value of the residues corresponding to the collection of the pairs of points under the hypothesis of an additional Gaussian noise, and any pair whereof the absolute value of the residue exceeds K times $\hat{\sigma}$ is tagged as an aberrant pair. One may advantageously fix the value of the variable K to 2.5. The least medial square calculation method used is a method conventionally known.

The rough homography H_g is finally obtained by least square regression calculation carried out over all the pairs judged as non- aberrant. It should be noted that the calculation of H_g may be fine-tuned further by iterating the process described previously, new matching pairs of points being obtained by applying the homography H_g to the model.
One may explain the calculation of the rough homography in the form of an algorithm with:
- p1 : corresponding in the picture to a point p of the theoretic ground resetted by *Hm* (average homography)
- p2 : contour point closest to p1

1. For each point p belonging to the theoretic contour:
   a. p1=Hm(p)
   b. if p is a point belonging to a vertical line p2=p1+Gv(p1)
   c. if not p2= p1+Gh(p1)
2. Robust calculation of the homography on the basis of the collection of the couples (p1,p2) found
   a. Perform n random draws of 4 couples of points
   b. For each draw:
      i. Calculate linearly the homography on the basis of the 4 couples
      ii. Calculate the medial error

        a. For the homography having given the minimal medial error

            i. Keep the non-aberrant couples (those whereof the absolute value of the residue is smaller than K times $\hat{\sigma}$).

            ii. Recalculate the rough homography Hg on the basis of all these couples

c) the fine resetting

The previous phase has therefore enabled to generate a matrix of rough homography H_g which is close to the final solution. The present step consists in fine-tuning the parameters of this homography in order to produce a fine homography function H_f so that the model adjusted by said function is closest to the lines in the picture. To do so, a minimisation method of a function of several variables is implemented. This function is derived from the criterion $D(I,H)$ defined previously and one seeks the matrix H_f solution of the following minimisation:

$$H_f = \min_{H_i \in H} (\oint_T d_c(I,H_i,s).ds) \tag{1}$$

The function d_c can be broken down as the sum of two components, a vertical one and a horizontal one:

$$d_c(I,H,s) = I_v(s).d_v(I,H.s) + (1 - I_v(s)).d_h(I,H.s)$$

where $\begin{cases} I_v(s) = 1 \text{ if S belongs to a vertical ligne} \\ 0 \text{ if not} \end{cases}$

The function $d_v(I,p)$ (respectively $d_h(I,p)$) represents the absolute value of the value at the point p of the card of the vertical (respectively horizontal) distances calculated at the rough resetting phase.

The integral contained in the formula (1) is sampled using a Bresenham line travelling algorithm so as to process

only integer coordinates. At the end of the day, the quantity to be minimised can be written as follows:

$$H_f = \min_{H_i \in H} \left(\sum_{p \in B(T,H_i)} (I_v(p).d_v(l,p) + (1 - I_v(p)).d_h(l,p)) \right) \qquad (2)$$

5   with $B(T,H_i)$ representing the collection of the pixels (with integer values) belonging to the contour of the ground T resetted by the homography H_i.

By reason of the very high non-linearity of the function (2) to be minimised, direct calculation is not possible and one

10   prefers to use an iterative minimisation function of a function with several variables which are here the 8 parameters of the homography H_f to be assessed. Several techniques are possible and notably, statistic methods and determinist methods.

15   Statistic methods advantageously guarantee the convergence towards the global minimum of the function to be minimised. The related calculation cost is however prohibitive in most applications. Among these usable methods, one may mention the simulated annealing method

20   whereof an implementation may be found, if needed, in « Numerical Recipes in C », P412, The Art of Scientific Computing, Cambridge University Press 2001.

The determinist methods, although convergent, do not ensure final provision of the global minimum of the function.

25   The minimum obtained after convergence is but a local minimum which is often quite close to the initial value wherewith the algorithm is initialised (i.e. the parameters of H_g in our case). However, thanks to the previous steps having enabled to obtain a homography matrix H_g which is close to

30   the final solution, this type of method may be applied with profit.

At first, the gradient of the function to be minimised is not available and, consequently, the techniques exploiting this information in order to ensure rapid convergence of a

35   determinist method algorithm are not applicable. Therefore, a

so-called Powell method is applied here. This method is described in detail in the book « Numerical Recipes in C », P412, The Art of Scientific Computing, Cambridge University Press 2001 to which may be referred. It is based upon a principle of alternate single-dimensional minimisations, the minimisation being carried out alternately over the 8 parameters of the homography.